Automated Data Interpretation

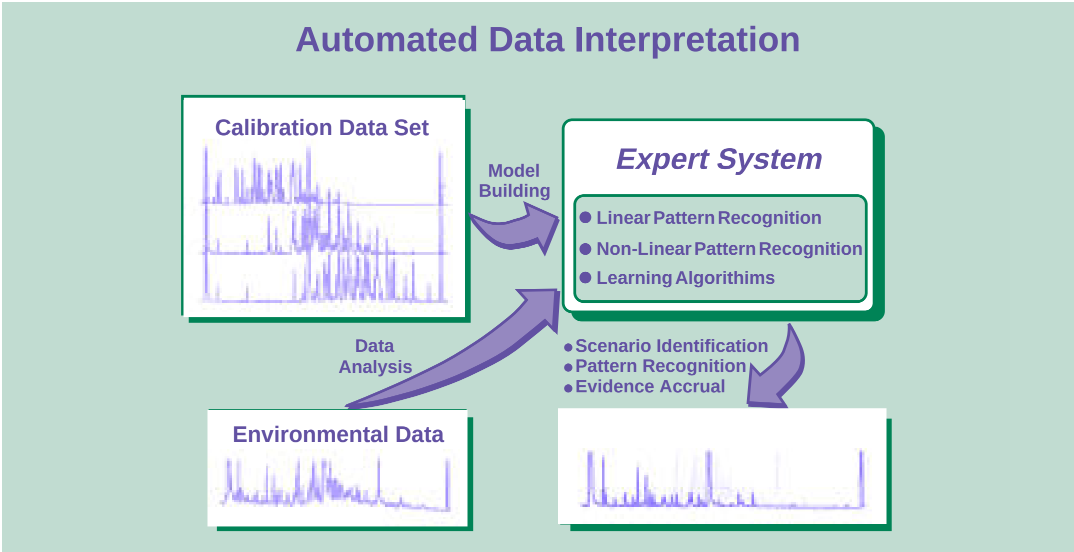

Figure 1. Automated analysis and automated data interpretation can be used to perform EPA Method 8080, the chromatographic analysis of environmental samples for multicomponent materials.

Task Description

The Contaminant Analysis Automation (CAA) Program is developing laboratory automation in a Standard Laboratory Module (SLM™) format. One of the goals of the CAA program is to link an automated analytical instrument module (AIM) and automated Data Interpretation Module (DIM) under the SLM™ format. The AIM is the element of the automated laboratory that performs the analysis on an aliquot of the prepared sample. This module will be a commercially available instrument that has been interfaced to the automated laboratory controller. The DIM is a software module in the automated laboratory that delivers chemical information about the sample from the raw data generated by the AIM.

The current target for the demonstration of automated analysis and automated data interpretation is EPA Method 8080, for organochlorine compounds and polychlorinated biphenyls (PCBs). Method 8080 is used to detect and quantitate Aroclors. Aroclors are the trade name for the mixtures of PCBs sold commercially. EPA method 8080 is targeted because chromatographic analysis of environmental samples for Aroclors, fuel spills, and other multicomponent materials are among the most difficult analyses to perform. The complexity of the chromatogram often inhibits the identification and quantitation of the sample components.

Pattern recognition tools will be used in the automated system to replace the intuitive pattern recognition used by an experienced chemist performing manual data interpretation.

These techniques will identify peak patterns from Aroclor PCB mixtures in the complex chromatograms of environmental samples. Standard pattern recognition methods will alleviate the inefficiency and lack of uniformity that result from subjective data interpretation decisions by a human expert. Current development includes principal component analysis, multiple linear regression, correlation, and neural network pattern recognition analysis methods.

The pattern recognition results will be interpreted in an expert-system-driven DIM, and expert-level rules and procedures will be incorporated into the expert-system DIM that will draw conclusions and provide confidence measurements from the data processing results. The figure illustrates the control of the data processing steps. The DIM will have two operational modes, an on-line mode that performs data analysis and is controlled by the automation controller via the SLM™ protocols, and an off-line mode that provides the analytical chemist with the tools required to build automated data analysis methods and review the raw data and analysis process. Other tasks required of the DIM include integrity validation of the raw data, the feedback of information to the task sequence controller to assure proper sample processing, automated quality control and quality assurance functions, and an interface to the laboratory database.

The data interpretation expert system is being developed in a standard format. Other analytical instruments will be incorporated into the automated laboratory and the techniques for analyzing their data integrated into the DIM. In addition to the GC, automated data analysis is being developed for GC / mass spectrometry instruments and atomic absorption spectroscopy instruments.

Technology Needs

There are several demonstrations of pattern recognition techniques applied to data analysis. However, no analytical methods have been developed for routine use in environmental laboratories. In the CAA Program, pattern recognition methods will be developed for data analysis and the applications will be tested and validated. The expert system integration of these techniques into routine methods will both enable automation of data analysis and provide a powerful aid to environmental chemists.

Accomplishments

- To test and demonstrate pattern-matching calibration model concepts, a typical calibration data set for the simultaneous analysis of Aroclors 1242, 1254 and 1260, according to EPA SW-846 Method 8080, were acquired. These Aroclors were selected because they were the primary mixtures sold and because state-of-the-art separations technology does not resolve all of the PCB structural isomers or congeners. This data set has been used to test principal component analysis pattern recognition.
- A principal component analysis of the Aroclor calibration data set was completed. The standard deviation of the Aroclor 1254 calibration function is 18 ppb. Essentially identical results were obtained for the other two Aroclors. The results clearly indicate that data interpretation techniques based upon pattern matching concepts can be implemented.
- Multiple linear regression, correlation, and neural network pattern recognition analyses have been applied to the Aroclor data. The expert system DIM draw upon the different strengths of the method suite to synthesize a final result.

Benefits

Currently, 40% to 60% of the cost of sample analysis is attributed to the effort of the data interpreation and data validation. Facilitating this effort with the automation being developed in the CAA Program will result in significant cost savings in sample analysis.

The instrumentation automation being developed in the CAA Program will make the current generation of laboratory testing instruments more robust. These instruments will be able to diagnose samples and instrument problems and take corrective action to ensure that subsequent samples are analyzed successfully. When incorporated into a fully automated laboratory, the instruments will be able to recommend additional sample preparation steps that are necessary to a successful analysis.

Collaboration/Technology Transfer

Participants in the data interpretation automation effort are as follows:

Academic collaborators:
Professor Mongi Abidi
University of Tennessee, Knoxville
Mark Krahling
University of Wisconsin - Eau Claire

Industrial collaborators:
Lockheed Missiles and Space Company
Varian Chromatography Systems Inc.
NeuralWare Inc.
Thru-Put Systems, Inc.

For further information, please contact: Mosiac Pages: http://eclipse.esa.lanl.gov/CAA and

John Elling
CAA Automated Data Interpretation
Los Alamos National Laboratory
(505)665-3047
elling@thermal.esa.lanl.gov

Bruce Erdal
Technical Program Manager
EM Program Office
Los Alamos National Laboratory
(505) 667-5338
erdal@lanl.gov

John Condren, Vice-President
SciBus Analytical, Inc.
(408)734-2100

Linton W. Yarbrough
DOE Headquarters Manager
Office Of Technology Development
(301) 903-7457
linton.yarbrough@em.doe.gov

SciBus Analytical, Inc.

Battelle Pacific Northwest Laboratories

INEL

Los Alamos National Laboratory

ornl

Sandia National Laboratories

University of Florida
University of Tennessee
University of Texas

LALP-94-33
April 1995

Los Alamos National Laboratory
Los Alamos, New Mexico 87545

A U.S. DEPARTMENT OF ENERGY LABORATORY